DOCUMENT RESUME

ED 247 058 RC 014 842

AUTHOR Gersten, Russell; And Others
TITLE Structured English Immersion for Hispanic Students in the U.S.: Findings from the Fourteen-Year Evaluation of the Uvalde, Texas, Program.
PUB DATE Apr 84
NOTE 34p.; Paper Presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 23-27, 1984).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; *Bilingual Education; Elementary Education; English (Second Language); *Hispanic Americans; *Immersion Programs; *Limited English Speaking; Longitudinal Studies; Program Evaluation; Scores; Spanish Speaking; Test Results
IDENTIFIERS *Project Follow Through; *Texas (Uvalde)

ABSTRACT

Evaluation of a structured immersion program, operating since 1968 in Uvalde, Texas as part of the United States Department of Education's Project Follow Through, suggests that the structured immersion approach has had a consistent positive effect on the academic achievement of the language minority students involved. Approximately 130 Spanish speaking Hispanic first grade students (83% classified as low income) enter the program each year, with approximately 100 students completing the 3 full years of the Follow Through program, which involves direct instruction by bilingual teachers and aides using a controlled vocabulary to teach needed skills and problem solving strategies to allow disadvantaged students to function at the same level as their peers. Evaluation of Comprehensive Test of Basic Skills, 1970 Metropolitan Achievement Tests, and Wide Range Achievement Test scores of students completing the program for 11 years indicate that achievement levels have been at or near grade level in mathematics and written language for over a decade. Performance on tests of oral reading are above grade level and scores in reading comprehension/vocabulary are at the 28th to 31st percentile, appreciably above typical levels for low income Hispanics. Follow-up studies conducted 2 and 3 years after students completed the program indicate significant, enduring effects in achievement. (NEC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED247058

Structured English Immersion for Hispanic Students in the U.S.: Findings from the Fourteen-Year Evaluation of the Uvalde, Texas, Program

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Russell Gersten
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
✓ This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

JUL 1984
RECEIVED
ERIC/CRESS

Russell Gersten
Robert Taylor[1]
John Woodward
W.A.T. White
University of Oregon

RC014842

Paper presented at AERA; New Orleans, LA, April 1984

[1] University of Arkansas at Monticello

Structured English Immersion for Hispanic Students in the U.S.: Findings from the Fourteen-Year Evaluation of the Uvalde, Texas, Program

In their comprehensive review of research on educational approaches for language-minority students in U.S. Schools, Baker and deKanter (1982) discuss the paucity of research on structured immersion programs in the U.S., despite the generally positive findings from these types of programs in Canada. A major reason little data is available on this approach is that federal policy in the 1970's made it difficult to implement such programs. This paper reports the results of an evaluation of a structured immersion program that has been in operation in Uvalde, Texas, as part of the U.S. Department of Education's Project Follow Through since 1968. Before discussing either the program characteristics or the evaluation findings, it will be helpful to contrast the structured immersion approach with others that are in use.

Structured English Immersion

Baker (1984) has identified three instructional alternatives for teaching students whose first language is not English: (1) Submersion, (2) Structured Immersion, and (3) Transitional Bilingual Education (TBE).

Submersion is a non-approach. Language minority students are merely placed in a regular classroom with no special help or program modifications to help them succeed. It has been aptly described as "sink or swim." The U.S. Supreme Court has ruled that submersion violates the civil rights of language minority students.

On a superficial level, structured immersion may seem similar to submersion. Most of the basic academic instruction in math and reading for the limited-English proficient (LEP) students is conducted in English. However, there is a critical difference. Though most new content is introduced in English, it is always introduced using vocabulary understood by the student. Teachers do not assume the children understand the English

language. Difficult new words are pretaught to the students, sometimes using the child's native language. The native language is rarely used by the teacher, unless it is necessary to help the student understand English language concepts. The philosophy for structured immersion is articulated by Baker (1984):

> The solution to developing English proficiency and progressing in other subjects is to teach all subjects in English at a level understood by the students. The curriculum assumes no prior knowledge of English. Language minority students in effect learn English as they learn math, and learn math through English instruction that is understandable at their level of English proficiency. In short, practice makes perfect, and English is best learned by using it as much as possible through the school day. (p. 2)

This was the model used in the primary grades in Uvalde since 1968.

Transitional Bilingual Educaton (TBE) is the prevalent form of bilingual education in use in this country. In TBE, students are also taught English, but receive most academic instruction (in reading, math, and language arts) in their native language until they master English.

The prevalence of TBE seems to be true mainly as a result of the Office of Civil Rights' response to the Supreme Court decision, the so-called Lau Remedies. The Office of Civil Rights actually used TBE as a standard for evaluating compliance with Title VII of the Elementary and Secondary Education Act of 1965. Another reason that TBE has seen wide use is that it seems "common sense" to use the child's native language for instruction until s/he masters English. Districts can point to classes being conducted in Spanish, Vietnamese, or Hmong as a clear and visible sign of their recognition of the language-minority children's needs and rights. It is much more

difficult to point to the underlying structure of an Immersion prgram. As a result of these and other factors, school districts, in an effort to comply with Title VII, have generally chosen to implement a program of the TBE type.

Unfortunately, the effectiveness of these transitional bilingual programs has not been clearly established. In their review of 39 research studies, selected from several hundred because of their methodological validity, Baker and deKanter (1982) concluded that TBE has had mixed success. Further, they found that a number of studies supported Immersion, a coordinated English-as-a-Second-Language (ESL) approach, and even Submersion, over TBE.

Research studies conducted on Structured Immersion approaches have been few in number. Of those studies available on Immersion, some have demonstrated impressive gains with middle class students in Canada. Genesee (1976) demonstrated immersion also was effective for low income Canadian students. Despite the success of immersion in Canada, many argue it will not work in the U.S. For example, Lambert (1984, cited in Baker, 1984) asserts

> The story is completely different for language minority young people. Immersion programs were not designed or meant for ethnolinguistic groups in North America that have some language other than English as the main language used in the home. To place such children in an initially all-English instructional program would be to misapply the immersion process in a harmful, subtractive way. Their personal identities, their early conceptual development, their chances of competing or succeeding in schools or in occupations, and their interest in trying to succeed would all be hampered by an immersion-in-English program. Fortunately, practical and valuable alternatives are now available to help these children...These transformations are not based on typical compensatory education or catch-up-in-English models... (p. 5)

The current paper discusses a program which was (and is) based on the principles of structured immersion and compensatory education, models which Tucker feels are inappropriate for low income Hispanic students. Our contention, based on the evaluation findings discussed below, is that this is a viable, effective model for language minority students. The next sections describe the context of the study, the nature of the educational programs, and an overview of the evaluation findings.

Background

In 1968, the Uvalde, Texas, school district jointed in Project Follow Through, a federal compensatory program for low SES children in grades 1-3. Uvalde is a small town halfway between San Antonio and the Mexican border at Piedras Negras. Although Uvalde is near some of the largest ranches in the country, much of the population has a low socioeconomic level.

Educational Model

Approximately 85 percent of the students in the program are eligible for free lunch (i.e., classified as low income). The students in the program are Hispanic; many classified as limited English proficient upon entry (ranging from 60 to 80 percent) are classified as LEP. Many students enter first grade speaking only Spanish (Gersten, 1981). Others have learned some English as a second language. Of the 130 students who enter the program each year, about 100 complete the three full years of Follow Through (Gersten, 1983). The national Follow Through Project consisted of twenty different educational models; each local project selected one of the model. Uvalde chose the Direct Instruction Model (Becker, Engelmann, Carnine, and Rhine, 1981). The Direct Instruction Model was implemented in 19 school districts throughout the U.S. with low income populations that included rural blacks, urban blacks, rural whites, Native Americans, and Hispanic students (Becker, 1977;

Becker & Gersten, 1982). Uvalde's project is unique in that its students were 98% Hispanic (Gersten, 1983).

The Direct Instruction Model, as implemented in Uvalde, is an example of structured immersion. DISTAR reading, language, and arithmetic programs are the backbone of the curriculum. DISTAR is a structured program using step-by-step instructions where explicit problem solving strategies modeled by teachers become implicit strategies used by students in their independent work. This approach allows immediate corrections of errors based on previous learning. Built into the programs are small group instruction, a high level of verbal interaction, and a large amount of distributed practice. Because of the language-minority population, oral language skills are more heavily emphasized in Uvalde than in other Direct Instruction projects. The Direct Instruction programs are carefully designed so that students will correctly interpret the concept(s) being taught with a minimum of confusion or misinterpretation. A carefully controlled vocabulary is used. Direct Instruction strives to teach needed skills and problem solving strategies in the most effective manner in the shortest period of time to allow disadvantaged students to function at the same level as their peers.

Another part of the Direct Instruction Model is the thorough monitoring of student progress through the administration of a Continuous Test of program objectives. Using test results and observations to assess effectiveness, on-site supervisors have worked closely with teachers and aides to improve their teaching. Also, outside consultants coordinated by a project manager supplemented local efforts, provided expertise, and assessed and revised the implementation.

Bilingual paraprofessional aides are used along with the teachers to teach the DISTAR programs in first and second grades. There is a limited

use of aides in grade 3. At Uvalde, from two thirds to three fourths of teahers are bilingual, allowing them to provide prompts in Spanish when necessary and expedient. Teachers of the classrooms with LEP students are always bilingual. A number of teachers originally began as aides in the project, completed teacher preparation programs, and have continued in Uvalde Follow Through as teachers.

All teachers of the classrooms with LEP students are always bilingual. All students are taught days of the week, months, counting, conversations, stories, and songs in both Spanish and English. Other bilingual activities include constructing bulletine boards that have a Spanish and an English half, providing directions in both languages if needed, and praising students in Spanish, then English. Over the years, a Spanish version of DISTAR Language and other bilingual oral language and reading programs have been used with many children, but they have not been a consistent component of the program. Though the amount has varied from year to year and teacher to teacher, the average student in the program probably experiences 10 to 30 minutes a day of instruction in Spanish. Spanish is used primarily as an integral part of the immersion approach.

EVALUATION OF THE UVALDE PROGRAM

Overview

Because this report summarizes results of several evaluation studies (Becker & Gersten, 1982; Becker & Engelmann, 1978; Gersten, 1981; Gersten, Carnine & Keating, 1984), a brief overview of the findings will be presented first, followed by descriptions of each study.

The children evaluated were the Uvalde students who attended Follow Through classes for three full years - first through third grade. Over 85 percent of these students' families qualify as low-income households under Follow Through guidelines. (In one study, only the low income students are included. The longitudinal resuearch includes all students.)

According to Danoff's (1978) evaluation of Title VII, students such as these (i.e. ones in transitional bilingual programs) tend to perform below the 20th percentile in reading and and the 30th percentile in math on standardized achievement tests. This kind of performance has been recently corroborated in Uvalde, where entering first grade students scored at the 21st percentile in 1980 and at the 19th percentile in 1981 on the pre-Reading section of the California Achievement Test (Level II).

The measures used in this study were the Comprehensive Test of Basic Skills (CTBS), the 1970 Metropolitan Achievement Tests (MAT) and the Wide Range Achievement Test (WRAT). Scores on the latter two tests form the basis of the claims presented below.

The MAT is one of the more valid standardized achievement tests (Buros, 1978). The WRAT Reading is a measure of Decoding only.

Uvalde is a small community and virtually all of the lowest income students and the majority of LEP students are in Follow Through. Thus, it is impossible to find a local comparison group with equivalent students. In one study, the fifth-sixth grade followup (Becker & Gersten, 1982), were we able to find a group of comparable students in neighboring communities. this was not feasible on the other occasions. Thus a range of quasi-experimental designs were used. Below is an overview of the findings.

The measure of the program's effectiveness with these students is in the weight of the data - its consistency across cohorts over 11 years. This is evident in study one, which details performance in the areas of language, math, and reading.

Overview of Findings: Study 1

1. At the end of the third grade, the Uvalde Follow Through students consistently achieved above or near the national norm on the Language subtest of the Metropolitan Achievement Test (MAT), a test of skills in using written English. This performance is significantly superior to the level of low income students nationally (as reported, for example, by Molitor et al, 1977). This level is significantly above levels for LEP low income Hispanics as reported by Danoff (1978).

2. Uvalde Follow Through students also perform at, near or

slightly above the national median level in MAT Total Math. This effect has been replicated over eight cohorts of children.

3. Uvalde Follow Through students consistently perform between the 26th and 31st percentiles in MAT Total Reading at the end of third grade. The MAT tests vocabulary (word knowledge) and comprehension. While the Reading scores are lower than the Math and Language scores, they are still above levels reported by Danoff (1978), who found low income Hispanic students in bilingual programs reading at or below the 20th percentile. This was reflected in scores on the reading subtest of the MAT.

4. They consistently made significant gains against the standardization sample of the Wide Range Achievement Test in Reading. WRAT Reading is a measure of word attack skills.

Stability of effects. These effects are replicated over 10 cohorts of children over a 12 year period, demonstrating clearly that the effects are attributable to the intervention (Cook & Campbell, 1979) and can be replicated with a variety of teachers, aides, and administrators.

Relationship between level of teaching and achievement. There is a moderate correlation between the level of implementation of Direct Instruction programs and the achievement gains in reading.

Longitudinal research. Longitudinal followup studies have demonstrated enduring effects of the program years after the students leave. There appear to be significant effects in most

domains of academic achievement two to three years after students leave the program (Becker & Gersten, 1982). More recent research, (Gersten, Carnine & Keating, 1994) demonstrates that the programs have helped diminish the dropout rates for this group of students up through the high school years and it has significantly reduced the number of retentions.

Study One: Norm Referenced Evaluation of Achievement

1. Table 1 reports the sample size, mean normal curve equivalent, and percentile equivalent (For cohorts 3,4,5 and 6 the mean standard score was converted to an NCE.) Cohorts 3 through 8 were tested on the Elementary form, 1970 version of the Metropolitan, cohorts 10 through 13 on the newer (1978) version. Note the sample sites are quite large, ranging from 84 to 110. Uvalde Follow Through students in the program for three years were above or near the national norm on MAT Language.

The MAT tests usage, tense, punctuation, and basic grammar. It assesses written English only.

All but three cohorts scored above the national norm and seven cohorts are above the 60th percentile. Such scores are far higher than expected in a disadvantaged bilingual community.

Table 1

Total MAT Language for 11 Cohorts of Students in Uvalde Follow Through (1970 -1983)

COHORT	BEGAN FIRST GRADE IN	N	MEAN NCE	PERCENTILE
3	1970	104	59.5	68
4	1971	89	43.5	74
5	1972	91	56	62
6	1973	84	69	82
7	1974	106	59.6	68.7
8	1975	120	54	58
9	1976	NA	NA	NA
10	1977	110	48.5	48
11	1978	103	53.8	57.6
12	1979	101	47.7	48.4
13	1980	97	47	45

2. On the MAT Total Math, (Table 2) the Uvalde students in the structured immersion program are consistently above the 30th percentile "standard" in math discussed by Danoff (1978). In fact, six of the cohorts scored at, near, or slightly above the national median level. Again, as Table 2 indicates, this level of performance is replicated over ten of the eleven cohorts, a clear indication that the effect was due to the intervention. Despite the year to year fluctuations, it is apparent the model is producing consistent educational effects. Note that the

Total Math composite include not only computation, but language-related skills in math problem solving and math concepts.

Table 2

MAT Total Math for Uvalde Follow Through Students (1970-1983)

COHORT	N	MEAN_NCE	PERCENTILE
3	103	51	52
4	89	56	62
5	91	53	56
6	84	53	56
7	106	44.5	40
8	120	44.5	40
9	96	44.5	40
10	110	47	45
11	102	47.6	46
12	102	39.7	31
13	97	43	37

3. Reading performance was assessed with two measures, the MAT and the WRAT. The MAT is a multiple choice test assessing word knowledge (vocabulary) and reading comprehension. The WRAT is an individually administered test of basic oral reading skills. It was used as a supplemental measure. Scores on the MAT are consistently above the 20th percentile, and in five cases they exceeded this mark by an average of of ten points. Table 3 presents the MAT data.

Table 3

MAT Total Reading for Uvalde Follow Through Students (1970-1983)

COHORT	N	MEAN NCE	PERCENTILE
3	105	39	30
4	89	36	26
5	91	39	30
6	84	39	30
7	106	36	26
8	120	33.5	22
9	96	33.5	22
10	110	36	26
11	104	38	29
12	102	40	32
13	97	41	34

The most dramatic effects of the model are evident in Table 4. The Direct Instruction reading programs (Distar) heavily emphasize decoding (word attack skills) at the beginning stages of reading.

In all cases, the students make significant gains against the standarization sample of the WRAT. (For the WRAT, a classical norm-referenced design, where students are tested on the same form and level of the subtest and then progress is compared to that of the norm sample used.) The reader should be alerted to the fact that the standardization sample of the WRAT

is small, and far from ideal (unlike the MAT, which is a carefully normed test, including a carefully stratified sample which includes language minority students). Nonetheless, the WRAT testing demonstrates significant gains in English language reading.

Table 4

WRAT Reading Gains of Follow Through Students Tested Entering First Grade and Leaving Third Grade: Norm-Referenced Comparison

		Pretest (Beginning of 1st Grade)			Post test (End of 3rd Grade)			Significance		
Cohort	N	SS	SD	Percentile	SS	SD	Percentile	t	p	°SD units
2	40	78.9	6.7	8th	112.9	18.0	81st	12.7	.001	2.75
3	68	79.6	6.1	9th	102.0	13.4	55th	16.2	.001	2.30
4	30	78.1	8.9	7th	103.9	15.0	61st	10.5	.001	2.16
5	80	78.0	6.2	7th	106.6	12.4	67th	21.6	.001	3.08

°Magnitude of effects in pooled SD units

Uvalde Follow Through students consistently make significant gains against the standardization sample of the Wide Range Achievement Test in Decoding. (Note that no claims are made for reading comprehension because of the bilingual program, and difficulties of assessing growth in this area for bilingual populations (see, for example, Becker, Gersten & Carnine, 1978).

Only those low-income Follow Through students who were tested at both times (pre-first and post-third grade) are considered, the gains in WRAT Reading are more dramatic. Table

4 provides data for paired t-tests and magnitude of effects measurements. The largest increase was by Cohort II from the 8th to the 81st percentile, a jump of 73 points. The average gain for all cohorts was 47 percentile points. Each cohort's gain is at least six times as large as the .25 to .33 SD unit gain usually considered to be educationally significant (see Tallmadge, 1977).

Evidence of the gains achieved by the Uvalde Follow Through students in decoding are seen as late as fifth and sixth grades (Becker & Engelmann, 1978; Becker & Gersten, 1980). Follow Through students in fifth and sixth grades scored at the 37th and the 42nd percentiles, respectively, on the average on WRAT Reading. Demographically similar Fifth- and sixth-graders in nearby communities who did not attend Follow Through had respective mean percentiles at only the 19th and 27th levels. The fifth-grade differences using analysis of covariance (with sex, income level and primary home language as covariates) indicated significant differences favoring Follow Through ($p < .001$ for both grade levels).

Exploratory Research: Linking level of implementation to growth in achievement

During the 1981-82 school year, all teachers were assessed on the Implementation Rating Form (Gersten, Meyer & Zoref, 1979) by the program director. The form examined aspects of teaching performance deemed central to effective teaching with this model - including immediate corrections of student errors, active involvement of all students in the group in all aspects of the

lesson, student success rate, systematic use of praise, and use of criterion-referenced test information to improve instruction. The internal consistency reliability of the form is .92 inter-rate reliability .87 (Gersten, Carnine, Zoref & Cronin, in press). Table 5 presents each teacher's implementation score and the gains in reading (in NCE units) for the class on the Metropolitan Achievement Test.

There is a moderate correlation between the level of implementation of the Direct Instruction model and yearly gains in reading achievement. The correlation for each grade level is, as follows:

Grade 1	.37
Grade 2	.18
Grade 3	.68

Except for second grade the relationship between quality of teaching is moderate.

Obviously, these findings are exploratory - in need of replications with larger sample sizes. But they do suggest, as does Tikunuff's (1983) findings that the quality of teaching within an immersion model (or a TBE model) must be measured, that factors for compensatory educational literature (e.g. Rosenshine, 1982) such as student success rate during lessons, procedures for correcting student errors, and clarity of teacher instructions are likely to have at least a moderate effect on achievement performance.

Longitudinal Research Studies of the Later Effects of the Uvalde Program

Fifth-Sixth Grade Study. In the spring of 1975, fifth grade students from the Uvalde Follow Through program were tested on all subtests of the Metropolitan Achievement Test (MAT) and level II of the Reading Section Intermediate Level of the Wide Range Achievement (WRAT). Similar testing was done again for sixth graders in 1976. This was done to determine the later effects of the Direct Instruction method on their academic achievement.

In Uvalde it was not possible to find an equivalent comparison group which were not in Follow Through, since the very poorest were in Follow Through. However, comparison groups were found in three communities within 20 miles of Uvalde which have students of similar backgrounds (La Pryor, Batesville, and Sabinal). The same battery of tests from the MAT and WRAT were given to the children in these groups.

Demographic information was collected on income level, sex, primary home language, number of siblings, and mother's education. These variable were used potential covariates in the analyses. Table 5 shows the levels on these variables between the Uvalde Follow Through (FT) and Non-Follow Through (NFT) local comparison group. As can be seen, the groups are fairly comparable except for slight differences in mother's education and proportion of income. These were corrected by analysis of covariance. On the average, few mothers in either group had more than an elementary school education.

We felt it was important to see if they maintained and built

on the progress they had made in the first 3 years of elementary school. The results of the study are especially interesting in the context of this report, as it was the only time when a local comparison group was used to help measure the relative effectiveness of the Direct Instruction Follow Through program in Uvalde.

Table 5

Comparison of Follow Through (FT) and Non-Follow Through (NFT) Groups in Uvalde on Selected Demographic Variables for the Fifth Grade

	NFT	FT
Sample Size	86	117
Mother's Ed.[a] Scale	2.34	2.86
Proportion Low Income	.95	.87
Proportion Non-Anglo	.95	1.00
Number of Siblings	6.24	6.09

[a] 5= High school graduates

Results Academic performance by the Uvalde Follow Through students on the WRAT reading and MAT subtests is shown in Table 6. Both statistical significance levels and magnitude of effects (in pooled standard deviation units) are reported. In considering magnitude of effect, .25 is regarded as

educationally significant (e.g. Becker & Engelmann, 1977). Strong, consistent effects are found in written language, reading (decoding and comprehension), spelling ($p < .10$), science ($p < .02$), and math problem solving. The effects are strongest on the WRAT reading ($p < .01$), which specifically tests decoding ability. This demonstrates that skills that children have truly mastered (such as English, word attack skills, etc.) do not diminish in the years after students have been taught by the Direct Instruction method. Math problem solving effects are consistently stronger than math computation. At first this would seem unusual for a program with a heavy emphasis on acquisition of basic skills. Yet the finding is consistent with the emphasis in the Distar arithmetic programs on teaching general-case problem-solving strategies, including basic algebraic principles. It appears that students were able to use the skills they learned in the primary grades to solve more advanced word problems. However, they were not learning the arithmetic computations typically taught in the intermediate grades (long division, borrowing) at the same rate.

Table 6

Mean Magnitude of Effects (in Pooled Standard Deviation Units) Follow-up of Uvalde Students who Began First Grade in 1970 (Cohort 3)

	5th Grade (in 1975)	6th Grade (in 1976)
N=	117	108
Test		
WRAT reading, Level II	.47	.52
MAT word knowledge	.03	.19
MAT reading (comprehension)	.21	.20
MAT language	.13	.39
MAT spelling	.01	.25
MAT math computation	.11	.04
MAT math concepts	.10	.31
MAT problem solving	.33	.23
MAT science	.15	.36

When considered in respect to findings from the other four sites in this of the fifth and sixth graders, we find that none of the outcomes signficantly favored the comparison groups at the .05 level, and 31 percent favored the Uvalde Follow Through group. Thus, it is extremely unlikely that these results were due to chance.

It is reasonable to conclude that if students learn skills and problem solving strategies well, they do not lose this knowledge. Furthermore, without effective instruction which

continues to build on these skills in the intermediate grades, the children are likely to lose ground against their middle income peers. In fact, limited English-speaking students such as the Uvalde students appear to lose the most.

Effects Up Through the High School Years

The data that follows is a summary of a follow up study comparing cohorts 1 and 2 to demographically similar students in Uvalde who began first grade in 1966 and 1967. Students from the two years prior to Follow Through 1966 and 1967 are labelled cohorts A and B and are compared with those from the first two years of Follow Through cohorts 1 and 2. Unlike the fifth and sixth grade study, an interrupted time series design is used. (A large number of students in all four of the annual classes failed to remain in the community for the entire 12 years. The percentages remaining in the community for the entire twelve years 1966 to 1969 respecetively, were 43.1%. 41.6%, 60.7%, and 44.1%. These figures are quite typical for longitudinal studies of this nature. The figures given in the results reflect data only the students who did reach high school for some period of time.)

A preliminary analysis of t he high school data shows that the Direct Instruction students:

-are more likely to receive a high school diploma,
-are less likely to be retained in any grade,
-show better attendance in 9th grade.

The results did not show appreciable differences between ninth grade Follow through and Non-Follow Through students in achievement scores or high school grades averages, but such differences are likely concealed by the differences in retention and dropout rates. If lower performing students are given an extra year to learn, they can make gains that would help to mask a Follow Through advantage.

Tables 7 and 8 show that a higher percentage of Follow Through students graduated than for the two preceding classes of students. Table 9 shows that there were fewer retentions among Follow Through students, and especially after the first year.

Finally, 1968 and 1969 Follow Through students had a higher proportion of acceptable attendance (less than 10 absences per year) while in high school than comparison students (68.9% and 83.3% respectively).

Table 7

Results of the Twelve Year Longitudinal Study

Cohort Comparison	Began First Grade	N	Percent Graduating from High School
A	1966	103	37.9
B	1967	97	42.3
Follow Through			
1	1968	87	59.8
2	1969	47	53.2

Table 8

Cohort Comparison	Began First Grade	N	Percent Dropout from High School
A	1966	103	60.1
B	1967	97	57.7
Follow Through			
1	1968	87	40.2
2	1969	47	46.8

Table 9

Cohort Comparison	Began First Grade	N	Percent Retained from High School
A	1966	110	43.2
B	1967	111	46.8
Follow Through			
1	1968	94	42.5
2	1969	62	23.0

Discussion

These findings strongly suggest that the structured immersion approach had a consistent, positive effect on the academic achievement of the language minority students in Uvalde who were involved. Achievement levels at the conclusion of the program have been at or near grade level in mathematics and writen language for over a decade. Performance on tests of oral reading are also above grade level. Scores in reading comprehension/vocabulary are at the 28th to 31st percentile, appreciably above typical levels for low income Hispanic students.

Follow up studies conducted two and three years after the students left the program indicate significant, enduring effects in achievement--particularly in the areas of oral reading, math problem solving, comprehension, and science. These data suggest that the skills the students learned to mastery during the primary grades (basic oral reading strategies from Distar reading, generalized strategies for solving mathematical problems, and to some extent, basic comprehension strategies) are retained. But students are not learning new skills (e.g., new computational skills, new vocabulary concepts) at an acceptable rate in a regular public school setting. This was true for Uvalde as well as the English-speaking sites (such as Dayton, OH) included in the Becker and Gersten study (1982). We would argue that some of the features of the Direct Instruction program, particularly those supported by the teacher effectiveness literature (e.g., Brophy & Good, in press) should be continued in the fourth through sixth grades. We also see a need for intensive work in English language vocabulary development, more than is currently provided by basal reader series (Becker, 1977; Beck, 1984). Nonetheless, the program, as currently implemented, appears to have enduring effects on the lives of its participants...at least the academic aspects of their lives. The later effects study conducted in high school demonstrated

that students in the program were (a) less likely to drop out of school, (b) less likely to be retained sometime during fourth through twelfth grades, and (c) more likely to attend high school regularly. (No effect was found for 11th grade achievement; however this effect may have been confounded by differential dropout and retention rates.) Noting the extremely high dropout rate for Hispanic students nationwide, these findings should be considered seriously.

These evaluations only looked at achievement in English. Future research should examine other aspects (including affective measures) and more carefully assess childrens' entry level and proficiency in Spanish as well as English. However, in terms of future success, English language achievement and high school graduation are certainly important determinants.

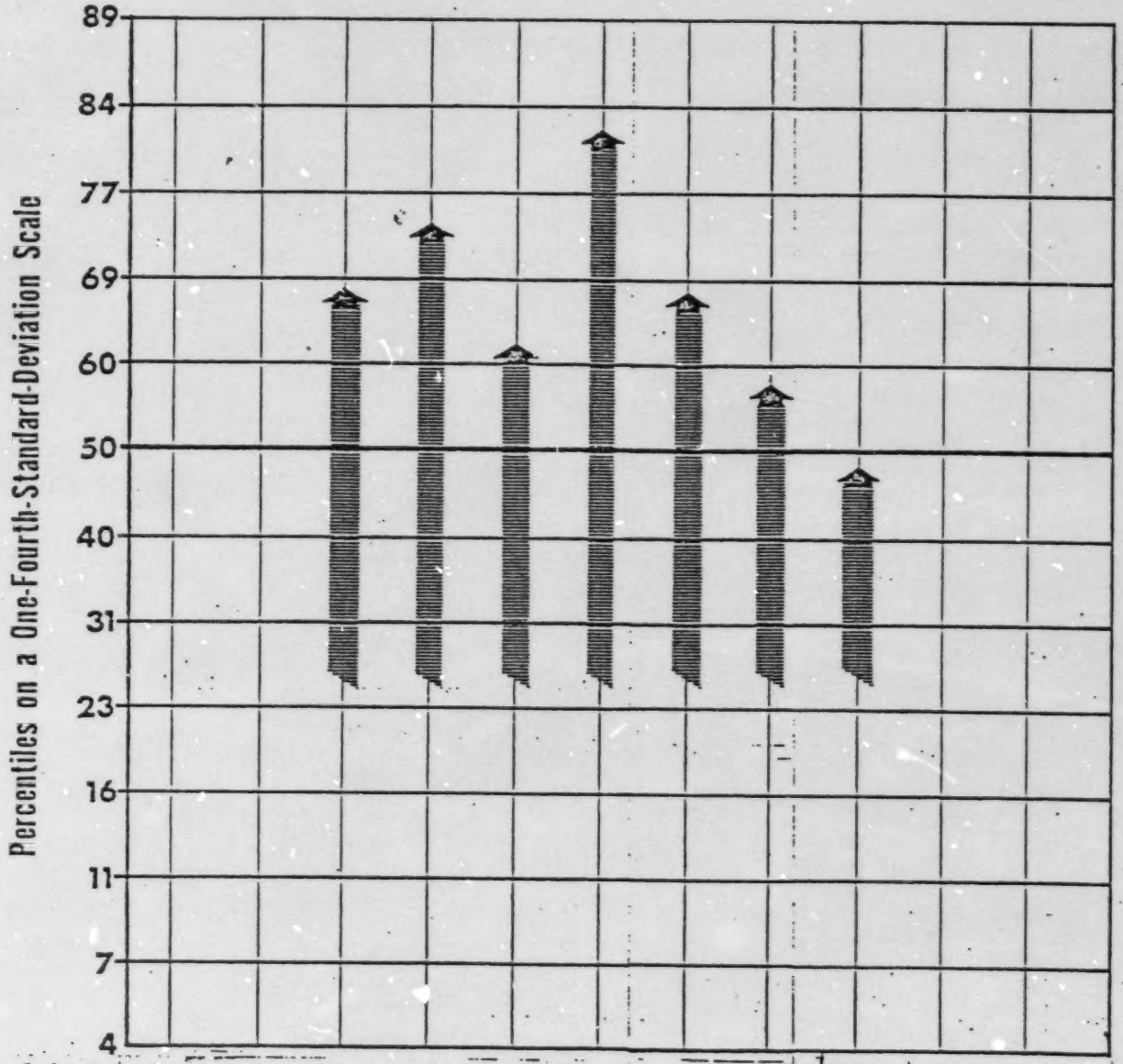
We strongly encourage further research on structured immersion programs in the U.S., but urge one caveat. It may be helpful for our understanding of bilingual education to be able to classify an educational program for language-minority students; the problem with a classification system is the range of characteristics of programs within each of the categories. If we are confronted with one structured immersion program that is successful and then another immersion program that is unsuccessful, the first question we must answer is: Do the two programs differ in essential instructional characteristics? To this end, the teacher effectiveness literature offers specific guidelines. Our exploratory study correlating achievement gains to teaching practices offers modest support for this approach. The work of Tikunoff (1983) and others on The Significant Instructional Features Study may be the beginnings of an attempt to really understand what constitutes effective instruction for language minority students. Both the structured

immersion and transitional bilingual education approaches are relatively easy to write about, and relatively difficult to implement sensitively on a day-to-day basis. Future research documenting exactly what happens in the effective classrooms, which aspects of the currently used curriculum are useful, and which could be modified, is an essential next step.

Stability of Effects

End of Third Grade Scores
(1971-1980)

M.A.T. Language
Low Income Students



Cohort	1	2	3	4	5	6	7	8	10
N	—	—	104	89	91	84	106	120	110
%-tile	—	—	68	74	62	82	68.7	58.0	48
S. S.	—	—	75.8	78.4	74.1	80.6	76.7	73.0	48.5[a]
S. D.	—	—	10.8	10.2	11.8	10.2	9.7	11.1	--

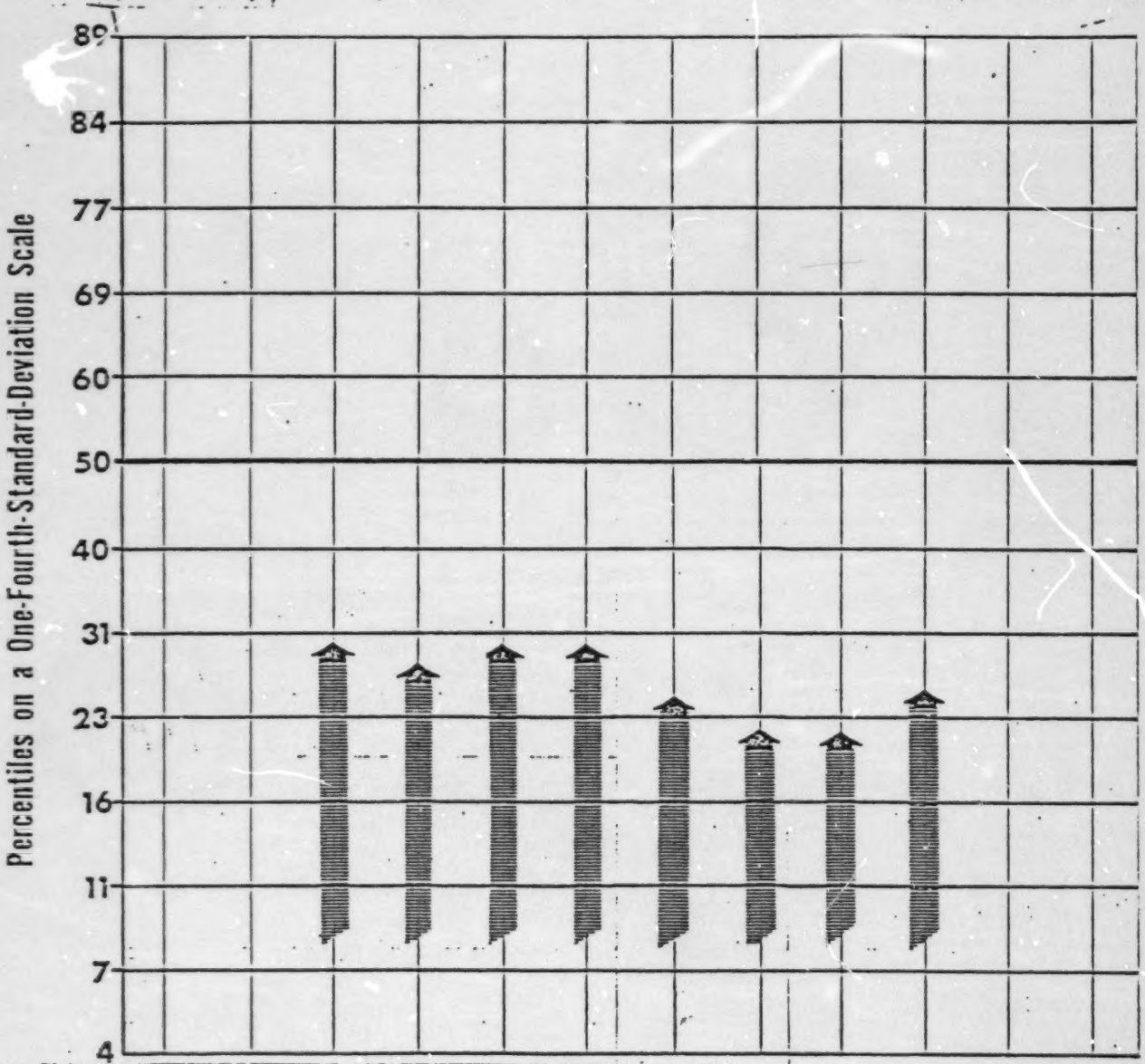
a NCE

BEST COPY AVAILABLE

Stability of Effects

End of Third Grade Scores
[1971-1980]

M.A.T. Reading
Low Income Students



Cohort	1	2	3	4	5	6	7	8	9	10
N	—	—	105	89	91	84	106	120	96	110
%-tile	—	—	30	28	30	30	25.8	22.1	22	26
S. S.	—	—	54.9	53.9	55.0	55.0	53.5	52.3	52	36[a]
S. D.	—	—	5.3	6.4	7.3	7.4	7.4	7.4	--	--

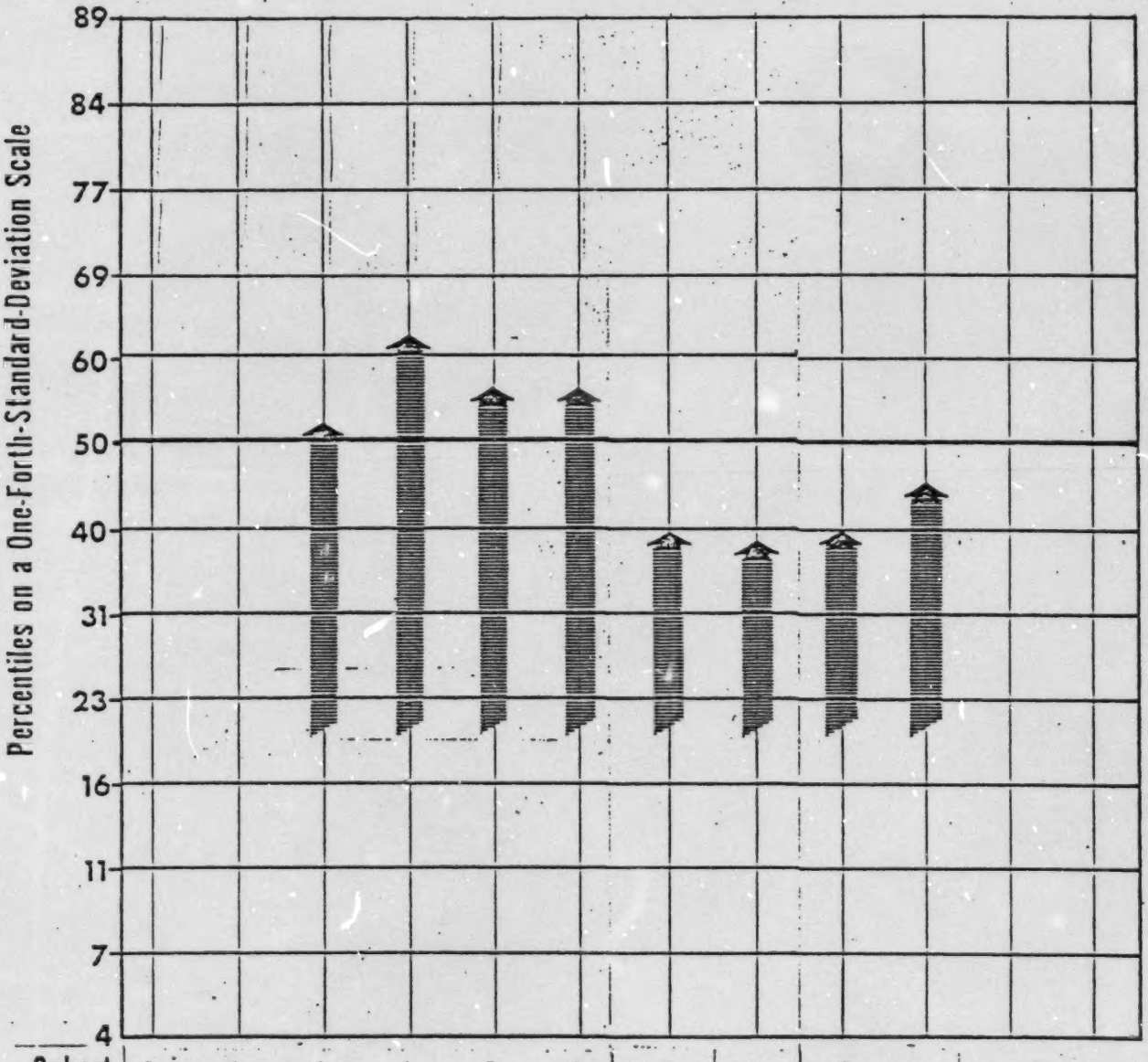
a NCE

BEST COPY AVAILABLE

Stability of Effects

End of Third Grade Scores
(1971-1980)

M.A.T. Math
Low Income Students



Cohort	1	2	3	4	5	6	7	8	9	10
N	—	—	103	89	91	84	106	120	96	110
%-tile	—	—	52	62	56	56	40.0	39.9	40	45
S. S.	—	—	70.9	73.7	71.7	72.4	68.1	68.0	68	47.2[a]
S. D.	—	—	7.7	8.9	9.7	9.5	9.2	9.1	--	--

a NCE

BEST COPY AVAILABLE

References

Baker, K. (1984, April). Immersion programs for language minority students: A review of the research and federal policy. Paper presented at AERA, New Orleans, LA.

Baker, K., & DeKanter, A. (1983). Federal policy and the effectiveness of bilingual education. Washington, D.C.: National Institute of Education.

Becker, W. C. (1977). Teaching reading and language to the disadvantaged--What we have learned from field research. Harvard Educational Review, 47, 518-543.

Becker, W. C., & Engelmann, S. (1978). Analysis of achievement data on six cohorts of low-income children from 20 school districts in the University of Oregon Direct Instruction Follow Through Model (Tech. Rep. #78-1). Eugene, OR: University of Oregon, Follow Through Project.

Becker, W. C., Engelmann, S., Carnine, D., & Rhine, R. (1981). The Direct Instruction model. In W. R. Rhine (Ed.). Encouraging change in America's schools: A decade of experimentation. New York: Academic Press.

Becker, W. C., & Gersten, R. M. (1979). Follow-up study of fifth and sixth graders: The 1976 replication study (Tech. Rep. #78-1). Eugene, OR: University of Oregon, Follow Through Project.

Becker, W. C., Gersten, R., & Carnine, D. (1979). Design and measurement issues in Follow Through evaluation research. Final report submitted to Office of Education, DHEW, pursuant to contract RFP 78-101.

Brophy, J., & Good, T. L. (in press). Teacher behavior and student achievement. In The third handbook of research on teaching.

Buros, O. (1978). Eighth mental measurement yearbook. Highland Park, NJ: Crypton Press, 1978.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis for field settings. Chicago: Rand McNally College Publishing Co.

Danoff, M. N. (1978). Evaluation of the impact of ESEA Title VII Spanish/English bilingual education program: Overview of study and findings. Palo Alto: American Institutes for Research.

Gersten, R. (1984, April). The effects of a structured immersion approach to minority language Asian students: Results of a longitudinal evaluation. Paper presented at annual meeting of the American Educational Research Association, New Orleans, LA.

Gersten, R. (1981). Final reports submitted to Joint Dissemination Review Panel (JDRP) of the U.S. Department of Education/NIE for Direct Instruction Projects in Williamsburg County, SC; New York, NY; East Las Vegas, NM; Uvalde, TX; Cherokee, NC. NIE/Department of Education.

Gersten, R., Carnine, C., & Keating, T. (1984, April). The lasting impact of the Direct Instruction Follow Through program: Preliminary findings of a longitudinal study of 1500 students. Paper presented at American Educational Research Association, New Orleans, LA.

Gersten, R., Carnine, D., Zoref, L., & Cronin, D. (1983). A study of educational change in an urban setting: A multi-level multi-instrument approach. Paper presented at American Educational Research Association, Montreal, PQ.

Gersten, R., Meyer, L., & Zoref, L. (1980). Implementation rating scale. Eugene, OR: University of Oregon, Follow Through Project.

Lambert, N. E., & Tucker, G. R. (1972). Bilingual education of children: The St. Lambert experience. Rawley, MA: Newbury House.

Molitor, J., Watkin, N., Napier, D., & Proper, E. C. (1977). Education as experimentation--The non-Follow Through study. Cambridge, MA: Abt Associates.

Rosenshine, B. (1983). Teaching function in instructional programs. Elementary School Journal, 83, 335-352.

Tallmadge, G. K. (1977). JDRP ideabook. Washington, D.C.: U.S. Government Printing Office.

Tikunoff, W. J. (1983). Significant bilingual instructional features descriptive study. San Francisco: Far West Labs.